



Archive-It Partner Meeting @ DLF
November 17, 2025

Blair-Caldwell African American Research Library
Denver, CO

Housekeeping



- **WIFI:** DPLWireless
- **Presenter slides:** will be shared with you by email.
- **Restrooms & water station:** Are at reception
- **Power outlets:** Are located on the floor
- **Masks:** Are available but not required. Please follow your own guidance and respect others' in regards to masking.

Code of Conduct

Internet Archive Code of Conduct: bit.ly/ia-coc

Internet Archive [events are] a friendly, safe and welcoming environment for all, regardless of sexual orientation, gender identity and expression, disability, personal appearance, body size, race, ethnicity, age, religion, nationality, or similar characteristics.



Welcome to Five Points



[A Brief History of the Blair-Caldwell Library](#)



[Denverite](#)

Agenda



- **9:30 AM:** Registration and refreshments
- **10:00 AM:** Welcome and introductions
- **10:30 AM:** Partners and Q&A
- **11:30 AM:** Internet Archive updates
- **11:50 AM:** Break
- **12:00 PM:** Focus groups and office hours
- **1:00 PM:** Lunch
- **1:30 PM:** Library and museum tour



Presentations



Piece of KAIC (Kansas Archive-It Consortium)

Crystal Hutchinson, Fort Hays State University - Tebo Library

“So you’ve inherited a web archive...”

Katherine Fisher, Emory University

Alex Kinnaman, Virginia Tech University Libraries

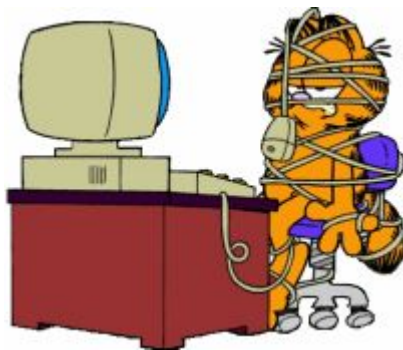
Laura Ruttum Senturia & Katie Rudolph, Denver Public Library

Internet Archive service and community updates

Kyrie Whitsett & Amanda McCabe, Internet Archive

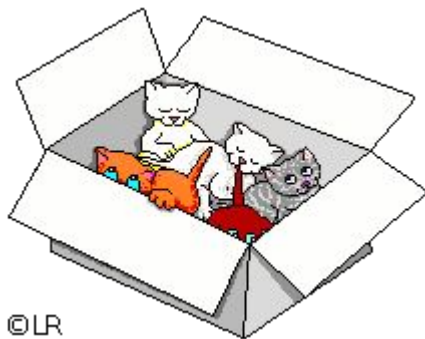
Archive-It Resources

Here when you need them:



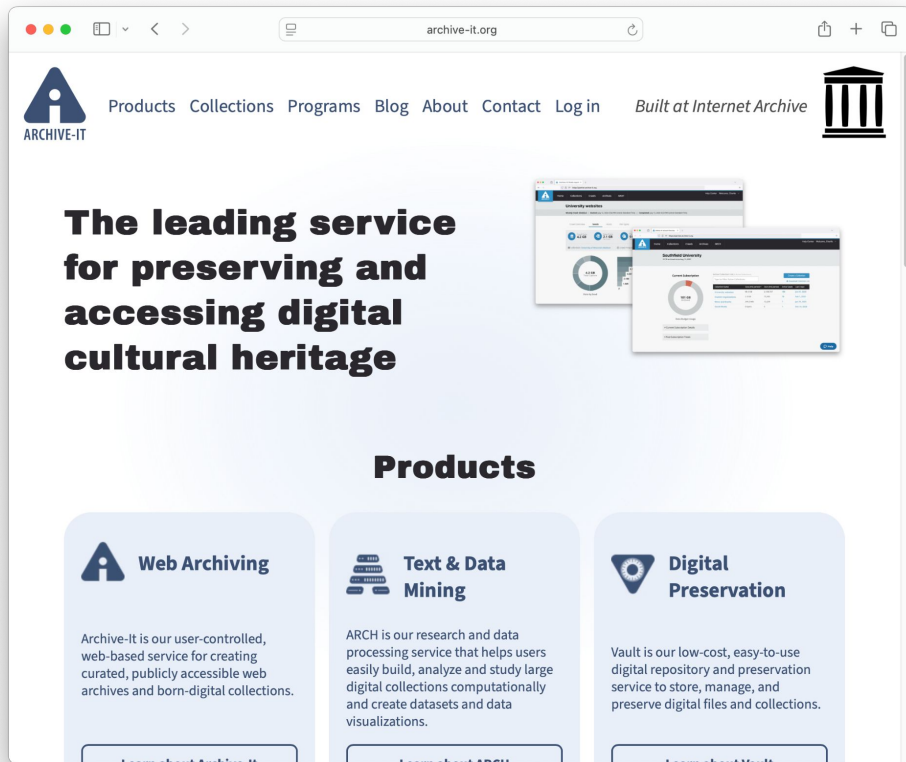
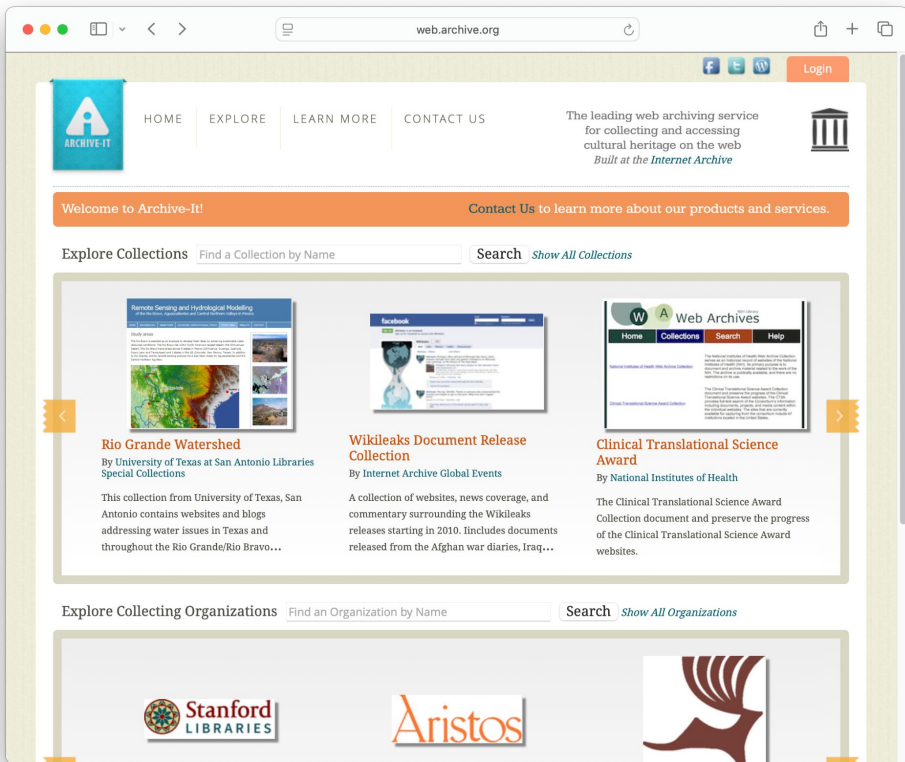
- [On-demand training](#) - Live Archive-It training from Internet Archive staff
- [Office Hours](#) - Real-time Q&A with a web archivist
- [Community Help](#) - Questions and answers from your Archive-It peers
- [Video Curriculum](#) - Short demonstrations of Archive-It tools and advanced training webinars

Recent updates:



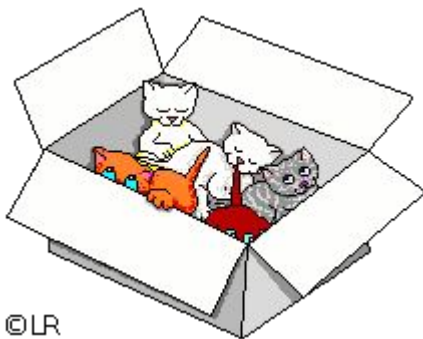
©LR

- **Public site redesign, Phase 1:**
 - Appearance, navigation, blog, & program pages
- **Crawling improvements**
 - PDF only for all crawlers
- **Smarter A/V handling**
 - Scope rules to exclude audio and videos
- **Bot protection/stability improvements**
 - Reduce disruptive automated traffic through new filtering and rate-limiting measures
 - Improved crawler access through collaboration with external services that manage site access controls and bot mitigation.



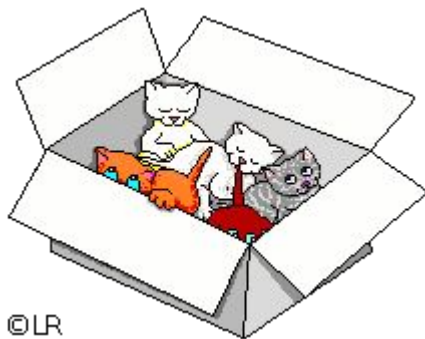
Coming up:

- **Smarter A/V handling**
 - Better deduplication of collected videos, starting with Youtube
- **Wayback Improvements**
 - Better, faster performance
 - Improved support for modern Javascript patterns
- **More efficient crawling and crawl scheduling**
 - Resumable Test crawls - coming soon!
 - Scheduling multiple crawls at any frequency, on different dates, and using different crawl tech (i.e. Standard and/or Brozzler)
- **Public site redesign, Phase 2**
 - Improve the search experience across collections
 - Support more flexible metadata display
 - Better highlight collection context and partner info



©LR

Themes shaping our development



©LR

- **Improving usability and accessibility** - Making Archive-It and web archives more intuitive for all types of users
- **Improving discoverability** - Exploring new ways to surface content within and across collections
- **Saving you time and building efficiency** - More task automation, reducing manual steps, and streamlining common tasks

We'd love your feedback

Resumable Test Crawls

Beta testers needed to...

- Help us validate this new workflow before release
- Try out in real scenarios and tell us what works (or what doesn't)

Email [**sylvie@archive.org**](mailto:sylvie@archive.org) to participate

Help Shape the Future of Search

We want to hear from you about...

- How search *should* work across your collections
- What you want to see in results (metadata? context? filters?)
- Your preferred search modes

Sign up for a call
[**https://archive-it.youcanbook.me/**](https://archive-it.youcanbook.me/)

Archiving & Data Services

Archives Research Compute Hub (ARCH):

- Released:
 - Speech, text, & (more) named entity recognition
 - Format-agnostic text extraction & expanded (Tr)OCR
 - archive.org, Vault & Wayback Machine support



Vault:

- Released:
 - Bulk metadata support
 - Growing API library
- Next:
 - Persistent Links
 - Bulk deposit application
 - Public access layer

Community Programs

Community Webs

- Program for public libraries and local history organizations
- Web archiving, digital preservation, and community archiving
- State cohort program
- Digitization projects



Collaborative ART Archive (CARTA)

- Web archiving program for art libraries and museums
- Engaging artists in the nomination process
- Raising awareness and promoting use of collections

Local News Preservation Program

- New program for newsrooms and organizations preserving local news
- National Summit on Local News Preservation
- Born Digital News Archive access portal

Discussion



Managing collections:

1. Introduce yourself and your relationship to web archiving
2. How/did you learn to manage web archive collections? What resources, docs, people, &c., did you rely upon?
3. What can Archive-It (the app or the community) do to support new web archivists?

Next generation search:

1. How search should work across your collections?
2. What you want to see in results (metadata? context? filters?)
3. Your preferred search modes?



Thank you!



Archive-It team

ait@archive.org

Karl Blumenthal, Senior Web Archivist

karlb@archive.org

Amanda McCabe, Product Manager, Vault

amccabe@archive.org

Kyrie Whitsett, Program Manager

kyrie@archive.org